

〈研究報告〉

マルチリンガルな意思伝達支援システムにおける 画像データベース構築のための調査研究

三 上 勝 生・山 本 裕 一

目次

- 1 はじめに
- 2 画像によるコミュニケーションの例とその限界
 2. 1 ピクトグラフ
 2. 2 アイコン
 2. 3 顔文字
 2. 4 手話
 2. 5 象形文字
- 3 画像解釈の自然言語依存性
 3. 1 画像の言語的体系化の不可能性
 3. 2 画像と自然言語の差異
 3. 3 画像記号の意味
- 4 対話的コミュニケーションにおける画像の利用価値について
 4. 1 状況および文脈依存文法
 4. 2 状況と文脈の分析
 4. 3 対話的コミュニケーションにおける画像の利用価値
- 5 画像データベース構築の方法論
 5. 1 画像データベース構築の限界と可能性
 5. 2 状況と文脈の「意味」
 5. 3 方法論と実践

1 はじめに

平成13年度からマルチリンガルな（多言語環境における）意思伝達支援システムの研究開発プロジェクト（通称「たまごプロジェクト」）

を立ち上げて、従来の機械翻訳システムとは原理的に異なる多言語間の同時翻訳を核とするコミュニケーション・システムの研究開発に取り組んできた。その自然言語間の「翻訳」とコミュニケーションに関する基礎研究と並行してわれわれが取り組んできたものに、自然言語によらない、画像を用いた意思伝達システムの研究がある。

平成14年2月の時点での具体的計画は、日常的な意思伝達に利用できる画像（動画を含む）の種類、厳密にはその概念的カテゴリーを特定し、それに見合う画像データを調査・収集、あるいは製作して、画像データベースを構築するという内容であった。また、その計画の意義については、この研究は学術的にはコミュニケーションに関わる多くの専門領域の研究にインパクトを与えるものであると同時に、その画像データベースは容易に各言語の音声データとも連携させうるものであり、将来的には、福祉分野をはじめ様々な分野での意思伝達ツールの開発に応用しうるものであると思われた⁽¹⁾。

しかしながら、様々な検討の過程で、当初想定していた「画像」の位置づけ、すなわちコミュニケーションにおける広義の「イメージ」の機能についての認識は、後述するように、深まると同時に変更を余儀なくされ、当初予定して

(1)平成14年度産業経営研究所研究調査申請書の事業内容より。

いた「画像データベース」の構築は根本から考え直す必要が生じた。そこでその再考の理由と現在検討中の新たな「画像データベース」の構築の方法論について、以下に概略を述べる。

2 画像によるコミュニケーションの例とその限界

2.1 ピクトグラフ

画像によるコミュニケーションの例としては、すでに道路標識や各種公共施設内の案内標識等に頻用されている、いわゆる「ピクトグラフ」がよく知られている。当初われわれはピクトグラフの果たしているコミュニケーションを人間同士の対話的コミュニケーションに拡張的に応用できるものと見込んでいた。しかしながら、以下に述べる理由から、それは不可能であるとの結論を得た。

ピクトグラフのメッセージ伝達の有効性は主にそれらが使用される状況と文脈が高度に限定されていることによる。その高度な限定は、画像の意味を強制的に学習させる社会環境、あるいは半強制的に学習させる生活環境の整備による。ピクトグラフによるメッセージ伝達の特徴は矢印に代表される方向と場所の指示、×印による特定行為の禁止命令等に典型的なように、限定された状況における一方的な指示と命令である。つまり、ピクトグラフとは基本的に特定の環境とそこを利用する人間との間のコミュニケーションのためのツールである。したがって、そのような基本的な特性に基づくピクトグラフを人間同士の対話的コミュニケーションに応用することはできない⁽²⁾。

2.2 アイコン

G U I (Graphical User Interface) の構成要素であるアイコンは極めて高度に限定された状況と文脈における画像を利用したコミュニケーションの例である。そこではコンピュータを使用するという限定された状況において、コン

ピュータに命令するという限定された文脈における使用者の操作を誘導するために画像が用いられる。しかしながらアイコンもまたコンピュータと人間の間のコミュニケーションを助けるツールではあっても、そのままでは人間どうしの対話的コミュニケーションをたすける手段にはなりえない。

2.3 顔文字

電子メールや掲示板、チャット等で文字と併用されるASCIIを駆使した顔文字は、ピクトグラフやアイコンとは異なって、人間同士の対話的コミュニケーションで一定の役割を果たしている。すなわち、顔文字は通常の文字情報だけでは表現しにくく伝達困難な感情的ニュアンスを付加的に表現し効果的に伝達する機能を担っている。

しかしながら仮にメッセージの内容を思想と感情に分けるならば、顔文字によって思想を表現し伝達することは不可能である。

2.4 手話

一般に手話は「視覚言語」に分類され、ある一定の範囲で自然言語からの自律性が認められる。しかし、実際にはその語彙表現法と文法形式は各国語、そしてさらには各国語内の諸方言の特性に依存する度合いを無視することはできない。したがって、われわれの観点からは各手話間の「翻訳」が問題になるため、手話を自然

(2)ここで、最初から単なる「標識」を超えて、人間同士のコミュニケーションに応用されるべく開発されたピクトグラフとして、世界的に有名な太田幸男による LCS (Lovers Communication System) を想起してもよい。しかしながら、その画像言語は既に習得している言語を手がかりにして、個々の画像の意味と構文構造を一から習得しなければならない「もう一つの言語」でしかない。そのような画像への「言語的アプローチ」は原理的にとりえないというのが本稿でのわれわれの思想的立場である。

言語間の翻訳に直ぐに役立てることはできない。

2. 5 象形文字

トンパ文字⁽³⁾のような象形文字について、あたかも自然言語なしで、トンパ文字を使用してコミュニケーションが可能であるかのような誤解がまかり通っているが、トンパ文字の有意義な使用はすでに中国語という自然言語の習得を暗黙の前提としたものにすぎない。

われわれが当初目指していたのは、状況と文脈への高度の依存を前提としない、あくまで人間どうしの対話的コミュニケーションに役立つような画像を活用したツールである。

3 画像解釈の自然言語依存性

3. 1 画像の言語的体系化の不可能性

したがって、われわれは従来の画像を用いたコミュニケーションを支える発想とは異なる角度から、画像一般がもつ本来的な特性を再考してみる必要に迫られた。

われわれの当初の目的は、記号としての画像のもつ多義性を曖昧さとして否定的にとらえるのではなく、柔軟さとして肯定的にとらえた上で、コミュニケーションを円滑に進行させる効果的なエージェント（代理）としての相対的に自律したシステムを「画像データベース」として構築することであった。すなわち、そこでは適度に体系化された画像群は自然言語と「並ぶ」もう一つの言語に近いものとして緩やかに想定されていた。したがって、「中間言語（普遍言

語）」⁽⁴⁾ カテゴリーが定まれば、他の自然言語とほぼ同等に交換されうる画像体系が構築可能であると考えられていた。

しかしながら、以下に述べるように、画像の本性からいっても、そのような画像の「言語的体系化」は不可能である。

3. 2 画像と自然言語の差異

メッセージの思想内容を画像で表現しようとする際に直面する最大の困難は、自然言語でおこなわれている世界の分節化、すなわちカテゴリー分類、のごく限られた範囲にしか画像は対応可能ではないということである。つまり、かなり焦点が絞られ、周囲から独立した対象物の描写にしか画像は使えない。これは画像の本来的な機能が常に「全体」を描写することによる。一方自然言語による思想表現では、「全体」は「部分」から構成される。すなわち文は語から構成され、構成されるということは分解できるということである。しかるに一枚の絵は部分に分解できない。したがって言語に画像を対応づけようとするれば、最初から単語に対応するような一枚の全体としての絵を用いるしかない。ところが、その場合でも絵はごく限られた対象物を描写することしかできない。しかも、例えば「リンゴ」、に対応するような絵を厳密には描くことはできない。絵は（写真でも）常に具体的な物を一定のアングルから一定のフレーム内で「全体的に」描写する。したがって、端的に「リンゴ」を表現することはできない。仮にある具体的なリンゴの絵を単語「リンゴ」として用いることを認めたとしても、例えば、動作や運動等のより複雑なカテゴリーを絵で描写する

(3)世界で唯一現在も使用されている象形文字。中国雲南省の納西（ナシ）族の間で用いられている。詳しくは下記のサイトを参照のこと。

「トンパ： www.templatebank.com/tompa/Default.htm」

「超漢字トンパサイト： www.chokanji.com/tompa」

(4)「たまごプロジェクト」(<http://www.sapporo-u.ac.jp/~tamago>) 参照。また、「マルチリンガル・コミュニケーション・システムの共同開発プロジェクト（たまごプロジェクト）の概要と経過報告」（札幌大学『経済と経営』第32巻第4号、2002年3月）参照。

際には、同種の困難がさらに深刻化する。

3. 3 画像記号の意味

画像を対話的コミュニケーションのツールの一部として利用できるであろうという期待は画像の本来的な多義性、あるいは無限の解釈可能性を過小評価した結果である。実際にある程度有効に機能しているように思われている画像的手段は実は自然言語による一義的解釈と了解を前提にしている。画像利用は確かにある一定の効率化には資するが、画像は本質的に対話的コミュニケーションにおけるメッセージの思想内容を構成する要素にはなりえない。画像をコミュニケーションの道具、一種の言語記号として利用しようとする企画は、実は言語的理解を前提にしてしか成り立たないものである。したがって、われわれは画像を一種の言語とみなす立場から離れて、状況と文脈を直接的に構成するものとみなす立場に移行せざるをえない。換言すれば、画像は状況と文脈を直示する、あるいは状況と文脈を描くものであるとみなす立場である。

4 対話的コミュニケーションにおける画像の利用価値について

4. 1 状況および文脈依存文法

機械翻訳が直面している原理的困難の一つは、メッセージの意味解釈の幅を絞り込む役割を果たす状況と文脈を、どのようにシステムに組み込むかということである。そしてわれわれの課題は人間が自然に行っている翻訳の過程にできるだけ則したシステムをいかに設計するかということである。従来のテキスト・レベルの機械翻訳が依拠する自由文脈文法では、状況と文脈に関する情報はどうしても「付随的」たらしざるをえなかった。しかしながら、その点でわれわれはいかなるメッセージもそもそも特定の状況と文脈におけるものであることを重視し、「状況および文脈依存文法」とも呼ぶべき立場を採

用する⁽⁵⁾。

4. 2 状況と文脈の分析

まずここで問題になる状況は当該システムが利用される状況によっておおまかに特定される。次にそのある程度特定された状況の中で生じるコミュニケーションの諸文脈が経験的に想定される。したがって、状況と文脈の具体的な特定は、どのようなシステムを開発するかに左右される。ところで状況と文脈が特定されるということは、それらは言語的に理解され言語表現されうること、そして状況と文脈をあらわす言葉も翻訳されてよいことを意味する。このことは、コミュニケーションを支援するシステムにおいて画像を利用する本質的理由は存在しないことを意味する。

4. 3 対話的コミュニケーションにおける画像の利用価値

それでは対話的コミュニケーションにおける画像の利用価値は全くないのであろうか。この点に関してもまた、実際にどのようなシステムを開発するかによって、そこで生じるコミュニケーションの特性によって、画像の利用価値もかなり変動することを認めなければならない。例えば、極端な例として、対話者の感情的諸要素が重視されるようなコミュニケーションが頻発するような状況で利用されるシステムの場合には、前述したようにいわゆる「顔文字」はかなり大きな効果を生む。また逆に事務的な手続きに近いやりとりにおいては画像の利用はかえって誤解や解釈の余地の拡大を招く恐れが大きくなる。

(5)周知のように、これはいわゆる「フレーム問題」につながる問題であるが、われわれは「人工知能」研究の文脈ではなく、あくまでわれわれのプロジェクトの内在的問題としてこの問題を定式化し探求する。

5 画像データベース構築のための方法論

5.1 画像データベース構築の限界と可能性

したがって、われわれは開発すべきシステムが実際に利用される状況と文脈があらかじめ特定されなければ、有効な画像データベースを構築することはできないことを認識しなければならない。これは画像データベース構築の限界を意味するが、しかし同時にその可能性をも示唆する。すなわち、状況と文脈を特定する機能を積極的に果たす役割を、画像に割り当てる道がある。

5.2 状況と文脈の「意味」

画像を言語記号の一種としてとらえるのではなく、むしろ状況と文脈そのものを表示するものとしてとらえることは先述した画像本来の特性にもかなった判断である。この場合、形式的には画像は言語的メッセージの背景をなす。しかし内容的には画像が表示する状況と文脈の「意味」によって言語的メッセージの意味もはじめて定まる。例えばGUIを成立させたデスクトップメタファーでは、机上空間を表示する画像構成が、人間からコンピュータへの命令（メッセージ）の意味を把握しやすくする。

5.3 方法論と実践

従来の自然言語処理においては、状況と文脈の「意味」はメッセージの言語的単位とみなされる文の意味からは切り離された上で、改めて事後的に付加する方法が探究されてきた。しかしながら、かりにわれわれの言語的コミュニケーションの「単位」が文であるとしても、その意味はそれがメッセージとして発せられる場合には字義通りの、辞書的な、あるいは語彙の積算的な意味ではありえず、必ずそれが使用される状況と文脈から意味の必須の部分を汲み上げている。したがって、自然言語処理研究においては、その目的、すなわちそれが具体的にどの

ような状況と文脈を想定するのかを可能な限り明確にしておかなければならない。これは画像データベース構築の観点から換言すれば、特定のコミュニケーションの舞台を構成する状況と文脈を表示する画像群をその特定の舞台装置のようなものとしてのデータベースとして構築する必要があるということである⁽⁶⁾。現在われわれは、その点を例証するための実験的なソフトウェアを企画し、それらの開発の準備段階に入っている。

*この報告は平成14年度産業経営研究所研究調査助成金に負うものである。

(6)ここでいう「舞台」あるいは「舞台装置」とは、いうまでもなく「インターフェイス」のことである。日常の対面状況とは異なり、コンピュータ端末を介したコミュニケーションが置かれた状況では、当該システムがどのようなインターフェイスをもつか、そこでのコミュニケーションの成否と質に決定的といってよい影響をもたらす。ここに、インターフェイス設計の重要さがある。